

## De CTEI: kwaliteit of kwantiteit of van allebei een beetje? Zorgen over een veel gebruikt meetinstrument voor de docentkwaliteit van de klinische docenten

### Inleiding

Binnen het medisch onderwijs wordt steeds meer aandacht besteed aan de kwaliteit van klinische docenten. De minister van VWS wil in het kader van de subsidieregeling van artsen in opleiding tot specialist (aios) de resultaten van 'tevredenheids-onderzoek' onder aios zelfs als parameter gebruiken bij de toewijzing van gelden. Het is dan ook belangrijk om bij de beoordeling van kwaliteit van docenten goede en gevalideerde onderzoeksinstrumenten te gebruiken. Eén van de vragenlijsten die daarvoor in toenemende mate wordt gebruikt is het (Cleveland) Clinical Teaching Effectiveness Instrument (de CTEI).<sup>1</sup> Deze lijst, ontwikkeld door Copeland en Hewson in de jaren negentig, bestaat uit 15 items en is voor het Engels taaldomein uitgebreid gevalideerd en betrouwbaar bevonden. In de Nederlandse situatie is door Van der Hem een betrouwbaarheidsonderzoek gedaan.<sup>2</sup> Sinds die tijd wordt de lijst ook in Nederland veel gebruikt.<sup>3-5</sup> De CTEI wordt voor zowel formatieve beoordelingen (evaluatie van docenten met het doel hen van feedback te voorzien; vaststellen van de kwaliteit van een groep docenten van een bepaalde afdeling) als voor summatieve beoordeling van docenten gebruikt.<sup>6-7</sup>

### Probleem

Bij het gebruik van deze lijst vermoedden wij een storende factor omdat dat de antwoordcategorieën in de gebruikte vijfpunts Likertschaal een combinatie vormen van een kwantitatief en kwalitatief oordeel. Antwoordmogelijkheden zijn bijvoorbeeld: 'always/superb' en 'never/poor'.

Inmiddels heeft Bierer in 2005 uitgebreid onderzoek gedaan naar het gebruik van de CTEI als summatief instrument met web-based invulmogelijkheid.<sup>6-7</sup> Zij gaat daarbij uit van het gegeven dat de lijst valide is. De antwoordcategorieën stelt zij echter niet ter discussie. Maar in een eerder artikel van Copeland en Hewson van 2000 zit al enige zorg: *'the largest source of error was due to trainees' interpreting items differently'*.<sup>1</sup> Een aanvullend, specifiek voor Nederland, probleem is dat zowel in de Nederlandstalige als Engelstalige publicaties van Nederlandse onderzoekers de letterlijke antwoordrubrieken niet eenduidig zijn (zie box 1). De dubbele antwoordmogelijkheid leidde bij ons tot grote twijfel over de interpretatie en daarmee over de validiteit van de uitkomsten van dit meetinstrument. Om hierover meer duidelijkheid te krijgen verrichtten wij een pilot-onderzoek.

### Vraagstelling

De vraagstelling van het pilot-onderzoek was of een aanpassing van de CTEI met onderscheid tussen kwalitatieve en kwantitatieve antwoordcategorieën duidelijk maakt dat de CTEI in de oorspronkelijke vorm niet eenduidig kan worden ingevuld.

### Methode

De antwoordcategorieën van de oorspronkelijke CTEI vragenlijst werden aangepast om de bovenstaande vraag te toetsen. Er werd een scheiding gemaakt tussen frequentie en kwaliteit van opleidersgedrag. In één vraag werd de frequentieaanduiding verwijderd ('geeft regelmatig feedback, zowel positief als negatief').

<b>Box 1</b> Verschillen tussen de diverse gebruikte versies antwoordcategorieën van de CTEI.		
Oorspronkelijke CTEI <sup>1</sup>	CTEI gebruikt in Nederlandse onderzoeken	
	Engelstalig <sup>3-5-8</sup>	Nederlandstalig <sup>9</sup>
1 = Never/poor	1 = Very poor/never	1 = Zeer slecht/nooit
2 = Seldom/mediocre	2 = Poor/seldom	2 = Slecht/zelden
3 = Sometimes/good	3 = Fair/sometimes	3 = Matig/soms
4 = Often/very good	4 = Good/often	4 = Goed/vaak
5 = Always/superb	5 = Very good/always	5 = Zeer goed/altijd
	1 = Never/poor	
	..... tussencategorieën	
	niet gespecificeerd	
	5 = Always/very good	

Aan de hand van deze aangepaste CTEI beoordeelden negen aiOS het docentgedrag van zes klinische docenten. De relatie tussen frequentie en kwaliteit van opleidergedrag werd statistisch getoetst (Wilcoxon-test).

### Resultaten

Alle scores beschouwend, was er overeenstemming bij 292 van de 509 scores (57%) in de twee aangepaste versies van de CTEI. Bij 43% was er echter een discrepantie tussen de scores van de twee versies: bij 153/509 (30%) scoorde de vraag hoger in de op kwaliteit aangepaste versie dan in de op frequentie aangepaste vraag. In 64/509 (13%) was dat andersom en scoorde de op frequentie aangepaste vraag hoger dan de op kwaliteit aangepaste vraag. Voor vijf van de 15 vragen was dat verschil tussen beide versies van de CTEI statistisch significant ( $p < .01$ ).

### Conclusies

Er is inderdaad sprake van een probleem bij het interpreteren van de CTEI. Degenen die de lijst gebruikten zijn logischerwijs afgegaan op het feit dat de lijst één van de weinige gevalideerde lijsten is op dit gebied. Mogelijkerwijs is bij het valideringsonderzoek aan de deelnemers niet de indeling

van de antwoordcategorieën voorgelegd. Wanneer de lijst gebruikt wordt als instrument voor het meten van kwaliteit van docenten waarbij op individueel niveau feedback aan docenten wordt gegeven, dan is het van groot belang of er door de invuller(s) op kwaliteit of op kwantiteit is gescoord. Een docent kan immers weinig frequent (seldom) uitstekend (superb) onderwijsgedrag vertonen en omgekeerd! Om de docent echter in staat te stellen zijn /haar gedrag te wijzigen, moet de feedback zo concreet mogelijk zijn.

We beseffen dat de steekproef bij ons pilot-onderzoek bijzonder klein is, maar toch lijkt dit steun te geven aan het door ons vermoede probleem. Het verdient daarom o.i. vooralsnog aanbeveling de lijst te gebruiken in zowel een frequentie- als een kwaliteitsvariant. Slechts dan is het mogelijk om concrete suggesties voor verbetering van het docentgedrag te formuleren op basis van de uitkomsten.

Daarnaast is een bijkomend probleem dat de verschillende auteurs ook afwijken van de oorspronkelijke antwoordcategorieën (zie box 1). Dit maakt het onderling vergelijken van uitkomsten van de verschillende onderzoeken riskant.

Tenslotte: de kwaliteit van clinical teachers gaat ons aan het hart. Wij willen

graag specifieke en concrete feedback aan onze docenten geven; de CTEI kan daarbij helpen mits de resultaten eenduidig te interpreteren zijn. Wij proberen met dit bericht een richting te wijzen voor een effectiever gebruik van de CTEI.

#### Literatuur

1. Copeland HL, Hewson MG. Developing and testing an instrument to measure the effectiveness of clinical teaching in an academic medical center. *Acad Med* 2000;75(2):161-6.
2. Hem-Stokroos HH van der, Vleuten CP van der, Daelmans HE, Haarman HJ, Scherpbier AJ. Reliability of the clinical teaching effectiveness instrument. *Med Educ* 2005;39(9):904-10.
3. Busari JO. The medical Resident as a Teacher. Teaching and Learning in the Clinical Workplace. [dissertatie] Maastricht: 2004; Universiteit Maastricht. [Dissertation; Maastricht: 2004; University of Maastricht].
4. Hem-Stokroos HH van der. The Clerkship as learning environment. [dissertatie] Amsterdam: 2005; VU. [Dissertation; Amsterdam: 2005; Free University of Amsterdam].
5. Bruijn M, Busari JO, Wolf BHM. Quality of clinical supervision as perceived by specialist registrars in a university and district teaching hospital. *Med Educ* 2006;40(10):1002-8.
6. Bierer SB. Psychometric properties of a clinical teaching effectiveness instrument used at the Cleveland Clinic Foundation. [dissertation] Ken State University; 2005.
7. Bierer SB, Hull AL. Examination of a Clinical Teaching Effectiveness Instrument Used for Summative Faculty Assessment. *Eval Health Prof* 2007;30(4):339-361.
8. Busari JO, Weggelaar NM, Knottnerus AC, Greidanus PM, Scherpbier AJJA. How medical residents perceive the quality of supervision provided by attending doctors in the clinical setting. *Med Educ* 2005;39(7):696-703.
9. Busari JO, Greidanus PM, Weggelaar NM, Schiering IAM. Supervisie in de medische praktijk: het oordeel van specialisten in opleiding. Arts-assistenten kindergeneeskunde beoordelen supervisors. *Tijdschrift voor Medisch Onderwijs* 2006; 25(4):170-9. [Supervision in the medical setting: residents in pediatric wards assess their supervisors. *Dutch Journal of Medical Education* 2006;25(4): 170-9].

P.M. Boendermaker en P. Remmelts

#### De auteurs:

P.M. Boendermaker, P. Remmelts zijn beiden senior staf-medewerker Wenckebach Instituut, Postgraduate School of Medicine, UMCG Groningen.